

(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION  
EN MATIÈRE DE BREVETS (PCT)

(19) Organisation Mondiale de la Propriété  
Intellectuelle  
Bureau international



(43) Date de la publication internationale  
15 septembre 2005 (15.09.2005)

PCT

(10) Numéro de publication internationale  
**WO 2005/086488 A1**

(51) Classification internationale des brevets<sup>7</sup> : **H04N 7/26**

(21) Numéro de la demande internationale :  
PCT/FR2005/050108

(22) Date de dépôt international :  
21 février 2005 (21.02.2005)

(25) Langue de dépôt : français

(26) Langue de publication : français

(30) Données relatives à la priorité :  
0450419 2 mars 2004 (02.03.2004) FR

(71) Déposant (pour tous les États désignés sauf US) : **THOM-  
SON Licensing S.A.** [FR/FR]; 46, quai Alphonse Le Gallo,  
F-92100 BOULOGNE (FR).

(72) Inventeurs; et

(75) Inventeurs/Déposants (pour US seulement) : **FRAN-  
COIS, Edouard** [FR/FR]; 18, allée du Locar, F-35890  
BOURG-DES-COMPTES (FR). **BOISSON, Guillaume**  
[FR/FR]; 12, rue Jean Malo-Renault, F-35000 RENNES  
(FR). **VIERON, Jérôme** [FR/FR]; 14, square Louis  
Massignon, F-35000 RENNES (FR). **MARQUANT,  
Gwenaëlle** [FR/FR]; 1, rue du Général de Gaulle, F-35340  
LIFFRE (FR). **ROBERT, Philippe** [FR/FR]; 7, allée du  
Bois Louët, F-35235 THORIGNE-FOUILLARD (FR).

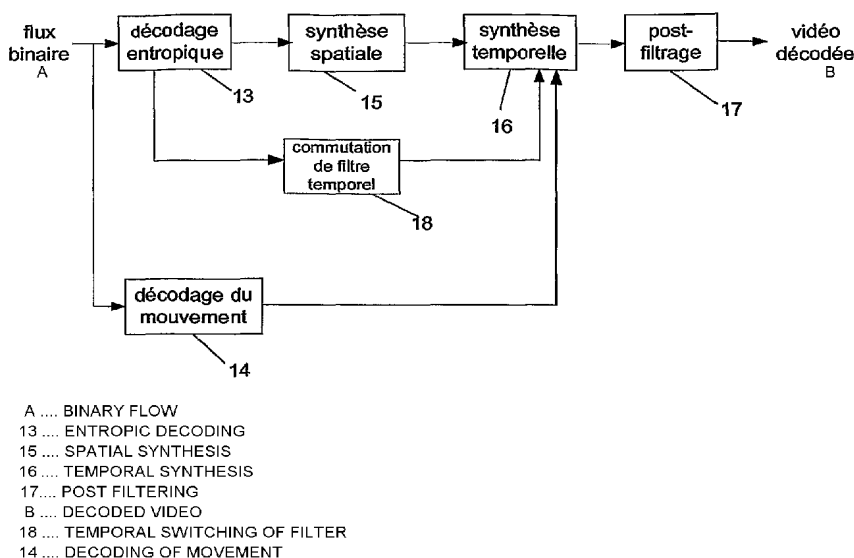
(74) Mandataire : **RUELLAN-LEMONNIER, Brigitte**;  
Thomson, European Patent Operations, 46, quai Alphonse  
Le Gallo, F-92648 BOULOGNE cedex (FR).

(81) États désignés (sauf indication contraire, pour tout titre de  
protection nationale disponible) : AE, AG, AL, AM, AT,  
AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO,

[Suite sur la page suivante]

(54) Title: METHOD FOR CODING AND DECODING AN IMAGE SEQUENCE ENCODED WITH SPATIAL AND TEMPO-  
RAL SCALABILITY

(54) Titre : PROCEDE DE CODAGE ET DE DECODAGE D'UNE SEQUENCE D'IMAGES CODEE AVEC ECHELONNABI-  
LITE SPATIALE ET TEMPORELLE



(57) Abstract: The invention relates to a method which is characterised in that the resolution selected for the movement information and the complexity of the interpolation filters used during a temporal filtering operation compensated in movement, depend on a decoding scenario, namely flow, temporal and spatial resolutions selected for the decoding either in terms of the corresponding temporal decomposition or of a combination of said parameters. The inventive method can be applied to so-called scalable video encoders/decoders, for example in the field of videotelephony or video transmission on the internet.

[Suite sur la page suivante]

WO 2005/086488 A1



CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) États désignés (sauf indication contraire, pour tout titre de protection régionale disponible) : ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), eurasién (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), européen (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO,

SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Publiée :**

- avec rapport de recherche internationale
- avant l'expiration du délai prévu pour la modification des revendications, sera republiée si des modifications sont reçues

En ce qui concerne les codes à deux lettres et autres abréviations, se référer aux "Notes explicatives relatives aux codes et abréviations" figurant au début de chaque numéro ordinaire de la Gazette du PCT.

(57) **Abrégé :** Le procédé est caractérisé en ce que la résolution choisie pour les informations de mouvement et la complexité des filtres d'interpolation utilisés, lors d'une opération de filtrage temporel compensé en mouvement, sont fonction d'un scénario de décodage, à savoir des résolutions spatiale, temporelle et du débit sélectionnés pour le décodage ou du niveau de décomposition temporelle correspondant ou d'une combinaison de ces paramètres. Les applications concernent les codeurs/décodeurs vidéo dits « scalables » par exemple dans le domaine de la vidéotéléphonie ou transmission vidéo sur internet.

**PROCEDE DE CODAGE ET DE DECODAGE D'UNE SEQUENCE D'IMAGES  
CODEE AVEC ECHELONNABILITE SPATIALE ET TEMPORELLE**

5

L'invention concerne un procédé de codage et de décodage vidéo d'une séquence d'image codée avec échelonnabilité spatiale et temporelle, par analyse temporelle hiérarchique exploitant le filtrage temporel compensé en mouvement.

10

Le domaine est celui de la compression vidéo basée sur des schémas à échelonnabilité spatiale et/ou temporelle dits aussi « scalables ». Il s'agit par exemple d'un codage en ondelettes 2D+t comportant un filtrage temporel compensé en mouvement.

15

Une chaîne de codage-extraction-décodage échelonnable est illustrée sur la figure 1.

Les images source sont transmises à un circuit de codage vidéo scalable 1. Le train binaire originel obtenu est traité par un extracteur 2 pour donner un train binaire extrait. Ce dernier est décodé par le circuit de décodage 3 qui fournit en sortie la vidéo décodée.

20

L'échelonnabilité ou « scalabilité » permet de générer un train binaire originel duquel on peut extraire des sous trains binaires adaptés à des ensembles donnés, débit, résolution spatiale, fréquence temporelle... Par exemple, si le train binaire scalable originel a été généré à partir d'une séquence vidéo à 25Hz, de résolution 720\*480 pixels et sans contrainte de débit, il est possible, après avoir extrait de ce train binaire les données adéquates, d'obtenir un sous-train binaire, par exemple de paramètres 1 Mb/s, 12,5Hz, 360\*240 pixels, lui-même scalable. Le décodage de ce sous-train binaire extrait génère une vidéo à 12,5Hz, de taille 360\*240 pixels.

25

Dans les approches existantes de compression vidéo scalables, le codage et le décodage procèdent de façon identique, sans prise en compte de conditions de fonctionnement telles que le niveau de décomposition temporelle, le débit, la résolution spatiale de la vidéo décodée... En particulier, si le décodage implique de la compensation de mouvement entre images, cette compensation est appliquée de façon identique, sans tenir compte de la taille des images, ni du débit de la vidéo à décoder. Cela résulte en une qualité d'images dégradée, en particulier lorsque la résolution d'images devient petite

35

par rapport à la taille des filtres d'interpolation utilisés pour la compensation de mouvement.

L'invention a pour but de pallier les inconvénients précités.

- Un des objets de l'invention est un procédé de décodage d'une
- 5 séquence d'images codée avec échelonnabilité spatiale et temporelle, les données codées comportant des informations de mouvement, comprenant une étape de synthèse temporelle hiérarchique effectuant un filtrage temporel compensé en mouvement, ou MCTF, d'images d'un niveau de décomposition en fréquence à partir des dites informations de mouvement, pour fournir des
- 10 images d'un niveau de décomposition inférieur, caractérisé en ce que, lors d'une opération de filtrage temporel compensé en mouvement, la résolution choisie pour l'exploitation des informations de mouvement et la complexité des filtres d'interpolation utilisés sont fonction d'un scénario de décodage, à savoir des résolutions spatiale, temporelle et du débit sélectionnés pour le décodage
- 15 ou bien du niveau de décomposition temporelle correspondant ou d'une combinaison de ces paramètres.

Selon une mise en œuvre particulière, le nombre de coefficients du filtre d'interpolation utilisé pour la compensation en mouvement dépend du scénario de décodage ou du niveau de décomposition temporelle.

- 20 Selon une mise en œuvre particulière, l'étape de synthèse temporelle hiérarchique est un décodage de coefficients ondelettes avec filtrage compensé en mouvement.

- L'invention concerne également un procédé de codage d'une
- 25 séquence d'images de résolution spatiale donnée, avec échelonnabilité spatiale et temporelle, comprenant une étape d'analyse temporelle hiérarchique effectuant un filtrage temporel compensé en mouvement, ou MCTF, d'images d'un niveau de décomposition en fréquence, à partir d'informations de mouvement entre ces images, pour donner des images d'un niveau de décomposition supérieur, caractérisé en ce que, lors d'une opération de filtrage
- 30 temporel compensé en mouvement, la résolution choisie pour l'exploitation des dites informations de mouvement et la complexité des filtres d'interpolation utilisés sont fonction de la dite résolution spatiale des images source ou du niveau de décomposition temporelle correspondant.

- Le procédé, selon une mise en œuvre particulière, comporte une
- 35 étape d'estimation de mouvement calculée entre deux images d'un niveau de décomposition donné pour réaliser la compensation de mouvement et en ce que la précision de calcul de l'estimation de mouvement est fonction du niveau

de décomposition temporelle ou de la dite résolution spatiale des images sources.

L'étape d'analyse temporelle hiérarchique est par exemple un codage en ondelettes avec filtrage compensé en mouvement.

5 L'invention concerne également un décodeur pour la mise en œuvre du procédé de décodage précédemment décrit, caractérisé en ce qu'il comporte un circuit de choix de configuration de mouvement pour déterminer la résolution du mouvement et le filtre d'interpolation à utiliser dans la compensation en mouvement pour le filtrage compensé en mouvement, en  
10 fonction du scénario de décodage, à savoir des résolutions spatiale et temporelles et du débit sélectionnés pour le décodage ou du niveau de décomposition temporelle correspondant ou d'une combinaison de ces paramètres.

15 L'invention concerne également un codeur pour la mise en œuvre du procédé de codage précédemment décrit, caractérisé en ce qu'il comporte un circuit de choix de configuration de mouvement pour déterminer le filtre d'interpolation à utiliser par le circuit d'analyse temporelle pour la compensation de mouvement en fonction de la dite résolution spatiale des images source ou du niveau de décomposition temporelle correspondant.

20 Selon un mode de réalisation particulier, le codeur est caractérisé en ce qu'il comporte un circuit de choix de configuration de mouvement pour déterminer la précision du mouvement calculé par le circuit d'estimation de mouvement, en fonction de la dite résolution spatiale des images source ou du niveau de décomposition temporelle correspondant.

25 La précision du mouvement et les filtres d'interpolation utilisés pour la compensation de mouvement dans le processus de codage et de décodage sont adaptés en fonction de différents paramètres, tels que le niveau de décomposition temporelle sur lequel on procède. Ces filtres sont adaptés,  
30 pour le décodage, au débit du flux décodé, à la résolution spatiale ou temporelle de la vidéo décodée.

Grâce à cette compensation en mouvement adaptative, la qualité des images est améliorée, la complexité des traitements est réduite.

35 D'autres particularités et avantages de l'invention ressortiront mieux de la description suivante donnée à titre d'exemple non limitatif et faite en regard des figures annexées qui représentent :

- la figure 1, une chaîne de codage selon l'art antérieur,

- la figure 2, un schéma de codage simplifié,
- la figure 3, un filtrage temporel d'un GOP,
- la figure 4, un filtrage temporel sur deux images,
- la figure 5, un circuit de décodage,
- 5           - la figure 6, un organigramme de choix de configuration de mouvement,
- la figure 7, un deuxième organigramme de choix de configuration de mouvement.

10           On considère un schéma de codage/décodage basé ondelettes 2D+t, opérant une analyse/synthèse ondelettes le long des trajectoires de mouvement. Le système opère sur des groupes d'images ou GOPs.

L'architecture globale du codeur est décrite en figure 2.

15           Les images sources sont transmises à un circuit d'analyse temporelle 4 qui effectue une analyse temporelle compensée en mouvement ou MCTF, acronyme de l'anglais motion compensated temporal filtering, pour obtenir les différentes bandes temporelles fréquentielles. Les images sont transmises à un circuit d'estimation de mouvement 7 qui calcule les champs de mouvement. Ces champs sont transmis à un circuit de « pruning » 10 qui effectue un « élagage » ou une simplification des informations de mouvement calculées par le circuit d'estimation de mouvement pour contrôler le coût du mouvement. Les champs de mouvement ainsi simplifiés sont transmis au circuit d'analyse temporelle afin de définir les filtres d'analyse. Ils sont aussi transmis à un circuit de codage 11 qui effectue le codage des champs de mouvement simplifiés.

25           Les images résultantes de l'analyse temporelle sont transmises à un circuit d'analyse spatiale 5 qui réalise un codage sous-bande de l'image bande basse et des images bandes hautes obtenues par l'analyse temporelle. Les coefficients ondelettes spatio-temporels ainsi obtenus sont finalement codés par un codeur entropique 6. Ce codeur fournit en sortie un ensemble de paquets binaires correspondant aux couches de scalabilités superposées, à la fois en qualité, en résolutions spatiales et temporelles. Un paquetiseur 12 effectue la fusion de ces paquets binaires avec les données de mouvement provenant du circuit de codage 11 pour fournir le train binaire échelonné final.

35           Les images aux différents niveaux de décomposition temporelle sont transmises par le circuit d'analyse temporelle 4 vers le circuit d'estimation

de mouvement 7 comportant un premier circuit de choix de configuration de mouvement. Ce circuit, non représenté sur la figure, définit les conditions de fonctionnement du circuit d'estimation de mouvement en fonction des différents niveaux de décomposition des images. De manière optionnelle, les informations de mouvement, une fois simplifiées via le circuit de pruning 10, sont transmises au circuit d'analyse temporelle à travers un circuit de commutation de mode 9. Ce circuit permet de tester la qualité de l'estimation de mouvement en testant par exemple le nombre de pixels connectés entre l'image courante et l'image précédente, à un niveau de décomposition donné, et peut imposer au circuit d'analyse temporelle un codage en mode intra ou un codage en mode prédictif, c'est à dire un filtrage de l'image courante avec l'image suivante et non précédente, lorsque cette qualité de mouvement est insuffisante. Le choix entre le mode intra et prédictif dépend par exemple de la qualité de l'estimation de mouvement entre l'image courante et l'image suivante. Le circuit d'analyse temporelle comporte un deuxième circuit de choix de configuration de mouvement, également non représenté sur la figure, qui détermine, en fonction des niveaux de décomposition des images et/ou de la résolution spatiale des images sources, la configuration à adopter pour la compensation de mouvement utilisée dans cette analyse temporelle.

20

La figure 3 représente de manière synthétique les opérations de filtrage temporel compensé en mouvement réalisées par le circuit d'analyse temporelle 4, avec une décomposition à 4 niveaux pour des GOPs comportant, dans cet exemple, 16 images représentées en traits gras.

25

Le mode de filtrage utilisé est appelé « lifting ». Au lieu d'utiliser un filtrage complexe pour le codage en ondelettes, exploitant un filtre linéaire de longueur importante, dans notre exemple le filtrage serait effectué sur un groupe de 16 images, ce procédé de filtrage consiste, de manière connue, à « factoriser » le filtre en exploitant des filtres de longueur limitée, par exemple deux si l'on choisit de filtrer les échantillons deux à deux, ce filtrage étant reconduit pour chaque niveau de décomposition. On considère donc le cas où le filtrage dans le sens du mouvement est effectué sur des paires d'images. Le filtrage basse fréquence et haute fréquence, sur chacune des paires du GOP, produit respectivement, au premier niveau de décomposition temporelle, 8 images de fréquence temporelle basse (t-L) et 8 images de fréquence temporelle haute (t-H).

35

Les images de fréquence temporelle basse sont ensuite décomposées à nouveau selon le même procédé. Le filtrage passe bas de ces images fournit 4 nouvelles images de fréquence temporelle basse t-LL et le filtrage passe-haut de ces mêmes images fournit 4 images de fréquence temporelle haute t-LH. Le troisième niveau de décomposition fournit 2 images de fréquence temporelle basse t-LLL et 2 images de fréquence temporelle haute t-LLH. Le quatrième et dernier niveau fournit une image de fréquence temporelle basse t-LLLL et une image de fréquence temporelle haute t-LLH.

Cette décomposition temporelle est une décomposition 5 bandes temporelles qui génère donc 1 image t-LLLL, 1 image t-LLH, 2 images t-LLH, 4 images t-LH, et 8 images t-H par GOP de 16 images. Les images t-L, t-LL, t-LLL et bien sûr les images originales sont ignorées pour le codage en aval puisqu'elles sont à l'origine de la décomposition en sous-bandes pour fournir les images décorréliées à chaque niveau. Cette décomposition permet donc une nouvelle répartition de l'énergie en générant une image utile de fréquence temporelle basse t-LLLL, qui représente une moyenne de l'ensemble du GOP et dans laquelle est concentrée l'énergie, et quatre niveaux d'images de fréquence temporelle haute de faible énergie, soit 5 bandes de fréquence. Ce sont ces images qui sont transmises au circuit d'analyse spatiale pour une décomposition spatiale en sous-bandes.

Pour réaliser les filtrages, un champ de mouvement est estimé entre chaque paire d'images à filtrer et cela pour chaque niveau. C'est la fonction de l'estimateur de mouvement 7.

Le filtrage d'une paire d'images source A et B consiste par défaut à générer une image basse fréquence temporelle L et une image haute fréquence temporelle H, selon les équations suivantes :

$$\begin{cases} L = (B + MC(A)) / \sqrt{2} \\ H = (A - MC(B)) / \sqrt{2} \end{cases}$$

où MC(I) correspond à l'image I compensée en mouvement.

La somme concerne le filtrage passe-bas, la différence, le filtrage passe-haut.

La figure 4 est une illustration simplifiée du filtrage temporel des deux images successives A et B, l'image A étant la première image selon l'axe des temps et selon l'ordre d'affichage, donnant une image basse fréquence L et une image haute fréquence H.



L'estimation de mouvement se fait par rapport à une image de référence, de l'image courante vers l'image de référence. On cherche pour chaque pixel de l'image courante, son correspondant, s'il existe, dans l'image de référence, et on lui affecte le vecteur mouvement correspondant. Le pixel de l'image de référence est alors dit connecté.

L'obtention de l'image L nécessite une compensation en mouvement de l'image A. Cette compensation se fait par estimation de mouvement de l'image B vers l'image A en prenant A comme image de référence, un mouvement et donc un vecteur étant alors affecté à chaque pixel de l'image B. La valeur d'un pixel de L est égale, au facteur de forme près, à la somme de la luminance du pixel correspondant de l'image B et de la luminance du pixel ou sous-pixel de A pointé par le vecteur mouvement attribué au pixel correspondant de l'image B. Une interpolation est nécessaire lorsque ce vecteur ne pointe pas sur un pixel de l'image A. Il s'agit ici de prédiction anticipée à partir d'une image de référence passée et d'un calcul de vecteurs mouvements anticipés ou « forward vectors » en anglais, en se référant à la norme MPEG.

L'obtention de l'image H nécessite une compensation en mouvement de l'image B. Cette compensation se fait par estimation de mouvement de l'image A vers l'image B en prenant B comme image de référence, un mouvement et donc un vecteur étant alors affecté à chaque pixel de l'image A. La valeur d'un pixel de H est égale, au facteur de forme près, à la différence de la luminance du pixel correspondant de l'image A et de la luminance du pixel ou sous-pixel de B pointé par le vecteur mouvement attribué au pixel correspondant de l'image A. Une interpolation est nécessaire lorsque ce vecteur ne pointe pas sur un pixel de l'image B. Il s'agit de prédiction différée à partir d'une image de référence postérieure et d'un calcul de vecteurs mouvements différés ou « backward vectors » en anglais, en se référant à la norme MPEG.

D'une manière pratique, seul un champ de vecteurs mouvement est calculé, de A vers B ou de B vers A. L'autre champ de vecteurs mouvement est déduit du premier, générant des pixels non connectés, c'est à dire non affectés d'un vecteur mouvement et correspondant à des trous dans le champ de vecteurs mouvement inverse.

D'une manière pratique, les images de fréquence basse et haute sont calculées comme suit:

$$\begin{cases} H = \frac{B - MC_{A \leftarrow B}(A)}{\sqrt{2}} \\ L = \sqrt{2} \cdot A + MC_{A \leftarrow B}^{-1}(H) \end{cases}$$

5 Ce filtrage, équivalent à celui précédemment décrit, consiste à calculer d'abord l'image H. Cette image est obtenue par différence point à point de l'image B et de l'image A compensée en mouvement. Ainsi, on retranche à un pixel de B une certaine valeur, interpolée le cas échéant, pointée par le vecteur déplacement dans A, vecteur mouvement calculé lors de l'estimation de mouvement de l'image B vers l'image A.

10 L'image L est ensuite déduite à partir de l'image H et non plus de l'image B, par addition de l'image A à l'image H compensée en mouvement inverse.  $MC_{A \leftarrow B}^{-1}(H)$  correspond à une « décompensation » en mouvement de l'image (H). Ainsi, on ajoute, à un pixel de A ou plus exactement à une valeur normalisée de la luminance du pixel, une certaine valeur, interpolée le cas échéant, située, dans l'image H, à la base d'un vecteur déplacement de B vers A et pointant le pixel de A.

15 Un même raisonnement peut être tenu au niveau d'un bloc d'image au lieu d'un pixel.

20 Le circuit d'estimation de mouvement 7 exploite par exemple un algorithme d'estimation de mouvement par appariement de blocs plus connu sous l'appellation anglaise de « block matching ». Un bloc d'image courant est corrélé aux blocs d'une fenêtre de recherche dans l'image de référence pour déterminer le vecteur de mouvement correspondant à la meilleure corrélation. Cette recherche est effectuée non seulement sur les blocs de la fenêtre de recherche obtenus par déplacements successifs d'un pixel en horizontal et vertical mais également sur des blocs interpolés si la précision recherchée est inférieure au pixel. Cette interpolation consiste à calculer les valeurs de luminance des sous-pixels, pour la génération de blocs d'images obtenus par déplacements successifs d'une valeur inférieure à la distance entre deux pixels. Par exemple pour une précision du quart de pixel, un test de corrélation est effectué tous les quarts de pixels, en horizontal et en vertical. Cette interpolation exploite des filtres appelés filtres d'interpolation de l'estimation de mouvement.

35 Les images pour lesquelles un filtrage temporel compensé en mouvement est à effectuer sont transmises à l'estimateur de mouvement 7 pour qu'il estime le mouvement entre deux images. Ce circuit comporte un premier

circuit de choix de configuration de mouvement qui reçoit, en plus de l'information de niveau de décomposition des images, d'autres informations telles que la résolution spatiale des images sources. Ce circuit décide de la configuration de mouvement en fonction de ce niveau et/ou de la résolution spatiale. Ainsi, par exemple, la précision dans le calcul des vecteurs mouvement dépend du niveau de décomposition temporelle des images traitées. Cette précision est d'autant plus faible que le niveau de décomposition est grand. Les filtres d'interpolation de l'estimateur de mouvement sont configurés pour être adaptés à cette précision de mouvement. Un exemple de configuration est donné par la suite.

Le circuit d'analyse temporelle 4, comme on l'a indiqué plus haut, réalise des compensations de mouvement pour le filtrage temporel des images. Ces opérations de compensations en mouvement nécessitent des opérations d'interpolation utilisant des filtres d'interpolation, et cela pour chaque niveau de décomposition. Le deuxième circuit de choix de configuration de mouvement, dans ce circuit d'analyse temporelle, qui peut être différent du premier, met en œuvre un algorithme de traitement adaptant la précision du mouvement et la complexité du filtre d'interpolation pour la compensation de mouvement en fonction du niveau de décomposition temporelle des images à compenser en mouvement. Comme pour le premier circuit de choix de configuration de mouvement, ces différentes adaptations ou configurations peuvent également dépendre de la résolution spatiale des images sources traitées.

Bien évidemment, un codeur ne comportant qu'un seul de ces circuits de choix de configuration est du domaine de l'invention.

Un décodeur selon l'invention est décrit en figure 5. Le flux binaire reçu par le décodeur est transmis à l'entrée d'un circuit de décodage entropique 13 qui réalise les opérations inverses du circuit de codage entropique du codeur. Il décode entre autres les coefficients ondelettes spatio-temporels et, le cas échéant, les modes de codage. Ce flux binaire est transmis en parallèle à l'entrée d'un circuit de décodage du mouvement 14 qui décode les champs de mouvement reçus dans le flux binaire pour les transmettre au circuit de synthèse temporelle. Le circuit de décodage entropique 13 est relié à un circuit de synthèse spatiale 15 qui reconstruit les images correspondant aux différentes sous-bandes temporelles. Les coefficients ondelettes temporels provenant du circuit de synthèse spatiale sont transmis à un circuit de synthèse temporelle 16 qui reconstruit les images de sortie à partir de filtres de synthèse

temporelle. Le circuit de synthèse temporelle comporte un circuit de choix de configuration de mouvement, non représenté sur la figure, qui détermine, en fonction des conditions du décodage et/ou des niveaux de décomposition des images, la configuration à adopter pour la compensation de mouvement utilisée dans cette synthèse temporelle. Le circuit de synthèse temporelle est relié à un circuit de post-traitement 17 dont la sortie est la sortie du décodeur. Il s'agit par exemple de post-filtrage permettant de réduire les artéfacts tels que les effets de bloc.

Dans le cas où le codeur exploite d'autres modes de codages que le mode MCTF, par exemple le mode intra et le mode prédictif, un circuit de commutation de filtre temporel est utilisé pour recevoir ces informations de mode de codage provenant du circuit de décodage entropique 13 et pour les transmettre au circuit de synthèse temporelle 16 qui effectue en conséquence les commutations de filtres.

Le circuit de choix de configuration de mouvement reçoit les informations de débit, de résolution spatiale et temporelle et les niveaux de décomposition temporelle. À partir de ces informations ou d'une de ces informations, il choisit, pour la synthèse temporelle, une configuration de compensation de mouvement. Le circuit de synthèse temporelle adapte le filtre d'interpolation en fonction de cette configuration choisie.

Le débit du flux binaire reçu par le décodeur correspond au train binaire extrait. Le codeur scalable transmet généralement le débit le plus élevé qui est le train binaire originel, comme on l'a vu précédemment, et l'extracteur, qui peut être commandé par le décodeur, extrait le train binaire correspondant aux résolutions souhaitées. Les informations de débit reçues sont disponibles au décodeur.

Les informations de résolution spatiale, temporelle et de débit définissent un scénario de décodage. Ce scénario est fonction par exemple de l'afficheur exploité par le décodeur, du débit disponible pour la réception des données. C'est à partir de ces informations et/ou du niveau de décomposition temporelle que le circuit de synthèse temporelle est configuré quant aux filtres d'interpolation.

Un exemple d'adaptation de la précision du mouvement et du filtre d'interpolation qui dépend de cette précision est donné ci-dessous, pour les opérations d'estimation de mouvement au codeur ou les opérations de compensation de mouvement au codeur ou décodeur :

configuration	précision du mouvement	filtres d'interpolation
1	1/4 de pixel	Bilinéaire
2	1/8 de pixel	¼ pixel par interpolation FIR 8 coefficients, puis 1/8 pixel par interpolation bilinéaire

Le filtre de la configuration 2 est très similaire de celui utilisé dans la norme MPEG-4 partie 10 (référence ITU-T Rec. H.264 ISO/IEC 14496-10 AVC).

5 La figure 6 représente un organigramme de décision mis en œuvre par le circuit de choix de configuration de mouvement appartenant au circuit d'analyse temporelle.

L'étape 20 détermine si la résolution de l'image source fournie au codeur est inférieure à celle du format QCIF, de l'anglais Quarter Common Intermediate Format et correspondant à 176 colonnes, 120 lignes. Dans l'affirmative, l'étape suivante est l'étape 23 qui décide la configuration 1.

Dans la négative, l'étape suivante est l'étape 21 qui vérifie le niveau de décomposition temporelle. Si ce niveau est strictement supérieur à 2, l'étape suivante est l'étape 23, la configuration 1 est choisie. Sinon, l'étape 15 suivante est l'étape 22 qui décide la configuration 2.

La figure 7 représente un organigramme de décision pour le décodeur.

L'étape 24 détermine si la résolution de l'image fournie par le décodeur et correspondant au flux binaire extrait est inférieure à celle du format 20 QCIF, 176 colonnes, 120 lignes. Dans l'affirmative, l'étape suivante est l'étape 26 qui choisit la configuration 1.

Dans la négative, l'étape suivante est l'étape 25 qui vérifie le niveau de décomposition temporelle. Si ce niveau est strictement supérieur à 2, l'étape suivante est l'étape 26, la configuration 1 est utilisée. Sinon, l'étape 25 suivante est l'étape 27. Cette étape 27 détermine si la résolution de l'image à décoder est égale à celle du format SD, de l'anglais Standard Definition, 720 colonnes, 480 lignes et si le débit du flux binaire reçu est inférieur à 1,5 Mb/s. Dans l'affirmative, l'étape suivante est l'étape 26 qui décide la configuration 1.

Dans la négative, l'étape 28 est l'étape suivante. Cette étape 28 30 détermine si la résolution de l'image à décoder est égale à celle du format CIF, 352 colonnes, 240 lignes et si le débit est inférieur à 700 kbits/s. Dans l'affirmative, l'étape suivante est l'étape 26 qui impose la configuration 1.

Dans la négative, la configuration 2 est imposée aux circuits de filtrage temporel.

Le filtre d'interpolation est par exemple du type FIR 8 coefficients, acronyme de l'anglais Finite Impulse Response. Le filtrage est effectué par convolution, en prenant donc en compte les luminances des 4 pixels précédents et suivants le sous-pixel à calculer.

Pour différentes positions au sous-pixel  $s$  à  $\frac{1}{4}$ ,  $\frac{1}{2}$ , et  $\frac{3}{4}$ , on peut utiliser trois filtres d'interpolation différents du type précédent. La valeur d'un coefficient  $n$  est donné par la formule :

$$f(n+s) = \sum_{m=-4}^4 h(m) \frac{\sin \pi (n+s-m)}{\pi (n+s-m)}, \quad 0 < s < 1.$$

$s$  est la position sous-pixel,  $s = \frac{1}{4}$ ,  $\frac{1}{2}$ , ou  $\frac{3}{4}$ ,  $n$  est le numéro du coefficient et  $h(m)$  le filtre d'atténuation ou fenêtre de Hamming.

On peut déduire des filtres FIR, par pondération par une fenêtre de Hamming et troncation de ces filtres pondérés.

Pour  $s = \frac{1}{4}$ , les coefficients sont :

[-0.0110 0.0452 -0.1437 0.8950 0.2777 -0.0812 0.0233 -0.0053]

Pour  $s = \frac{1}{2}$ , les coefficients sont :

[-0.0053 0.0233 -0.0812 0.2777 0.8950 -0.1437 0.0452 -0.0110]

Pour  $s = \frac{3}{4}$ , les coefficients sont :

[-0.0105 0.0465 -0.1525 0.6165 0.6165 -0.1525 0.0465 -0.0105]

Avec ces filtres, on peut interpoler au  $\frac{1}{4}$ ,  $\frac{1}{2}$  et  $\frac{3}{4}$  de pixel. L'interpolation se fait d'abord selon la dimension horizontale, puis verticale. L'interpolation au  $\frac{1}{8}$  de pixel est effectuée ensuite par une interpolation bilinéaire à partir des positions au  $\frac{1}{4}$  de pixel.

L'exemple d'adaptation donné ci-dessus au niveau du codeur peut s'appliquer de la même manière au niveau du décodeur.

D'une manière générale, le principe est d'utiliser une précision de mouvement limitée et des filtres d'interpolation simples quand on opère à des qualités d'image limitées, c'est-à-dire à bas débit, sur des images de petite taille et à des niveaux de décomposition temporelle élevés. A l'inverse, quand on traite des images de bonne qualité, résolution spatiale grande, débits élevés,

niveaux de décomposition temporelle bas, on utilise une grande précision de mouvement et des filtres d'interpolation sophistiqués. La justification de ce principe est que lorsque les images à filtrer sont pauvres en contenu fréquentiel ou de résolution limitée, il n'est pas utile d'employer des filtres d'interpolation très évolués ni une précision de mouvement très grande.

Les applications de l'invention concernent les codeurs/décodeurs vidéo dits « scalables » utilisés pour la compression/décompression de données, par exemple dans le domaine de la vidéotéléphonie ou transmission vidéo sur internet.

## REVENDICATIONS

1 Procédé de décodage d'une séquence d'images codée avec  
5 échelonnabilité spatiale et temporelle, les données codées comportant des informations de mouvement, comprenant une étape de synthèse temporelle hiérarchique (16) effectuant un filtrage temporel compensé en mouvement, ou MCTF, d'images d'un niveau de décomposition en fréquence à partir des dites informations de mouvement, pour fournir des images d'un niveau de  
10 décomposition inférieur, caractérisé en ce que, lors d'une opération de filtrage temporel compensé en mouvement, la résolution choisie pour l'exploitation des informations de mouvement et la complexité des filtres d'interpolation utilisés sont fonction d'un scénario de décodage, à savoir des résolutions spatiale, temporelle et du débit sélectionnés pour le décodage ou bien du niveau de  
15 décomposition temporelle correspondant ou d'une combinaison de ces paramètres.

2 Procédé selon la revendication 1, caractérisé en ce que le nombre de coefficients du filtre d'interpolation (16) utilisé pour la compensation  
20 en mouvement dépend du scénario de décodage ou du niveau de décomposition temporelle.

3 Procédé selon la revendication 1, caractérisé en ce que l'étape de synthèse temporelle hiérarchique (16) est un décodage de coefficients ondelettes avec filtrage compensé en mouvement.  
25

4 Procédé de codage d'une séquence d'images de résolution spatiale donnée, avec échelonnabilité spatiale et temporelle, comprenant une étape d'analyse temporelle hiérarchique (4) effectuant un filtrage temporel  
30 compensé en mouvement, ou MCTF, d'images d'un niveau de décomposition en fréquence, à partir d'informations de mouvement entre ces images (7), pour donner des images d'un niveau de décomposition supérieur, caractérisé en ce que, lors d'une opération de filtrage temporel compensé en mouvement (4), la résolution choisie pour l'exploitation des dites informations de mouvement et la  
35 complexité des filtres d'interpolation utilisés (9) sont fonction de la dite résolution spatiale des images source ou du niveau de décomposition temporelle correspondant.



5 Procédé selon la revendication 4, caractérisé en ce qu'il comporte une étape d'estimation de mouvement (7) calculée entre deux images d'un niveau de décomposition donné pour réaliser la compensation de mouvement (4) et en ce que la précision de calcul de l'estimation de mouvement (7) est fonction du niveau de décomposition temporelle ou de la dite résolution spatiale des images sources.

10 6 Procédé selon la revendication 4, caractérisé en ce que l'étape d'analyse temporelle hiérarchique (4) est un codage en ondelettes avec filtrage compensé en mouvement.

15 7 Décodeur pour la mise en œuvre du procédé selon la revendication 1, caractérisé en ce qu'il comporte un circuit de choix de configuration de mouvement (16) pour déterminer la résolution du mouvement et le filtre d'interpolation à utiliser dans la compensation en mouvement (16) pour le filtrage compensé en mouvement, en fonction du scénario de décodage, à savoir des résolutions spatiale et temporelles et du débit sélectionnés pour le décodage ou du niveau de décomposition temporelle correspondant ou d'une  
20 combinaison de ces paramètres

25 8 Codeur pour la mise en œuvre du procédé selon la revendication 4, caractérisé en ce qu'il comporte un circuit de choix de configuration de mouvement (4) pour déterminer le filtre d'interpolation à utiliser par le circuit d'analyse temporelle pour la compensation de mouvement (4) en fonction de la dite résolution spatiale des images source ou du niveau de décomposition temporelle correspondant.

30 9 Codeur pour la mise en œuvre du procédé selon la revendication 4, caractérisé en ce qu'il comporte un circuit de choix de configuration de mouvement (7) pour déterminer la précision du mouvement calculé par le circuit d'estimation de mouvement (7), en fonction de la dite résolution spatiale des images source ou du niveau de décomposition temporelle correspondant.

35

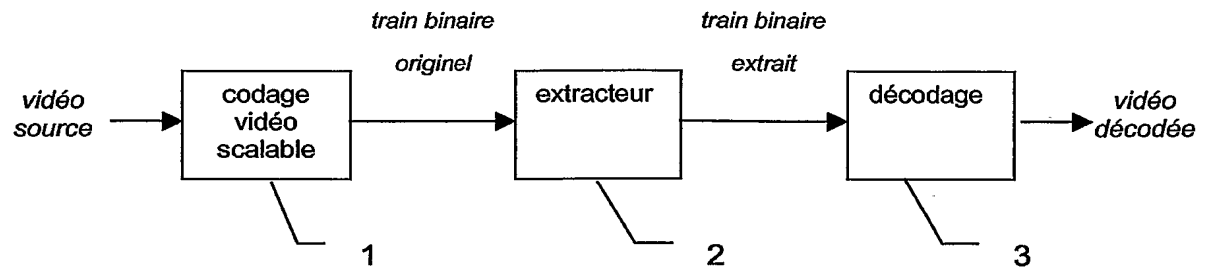


FIG.1

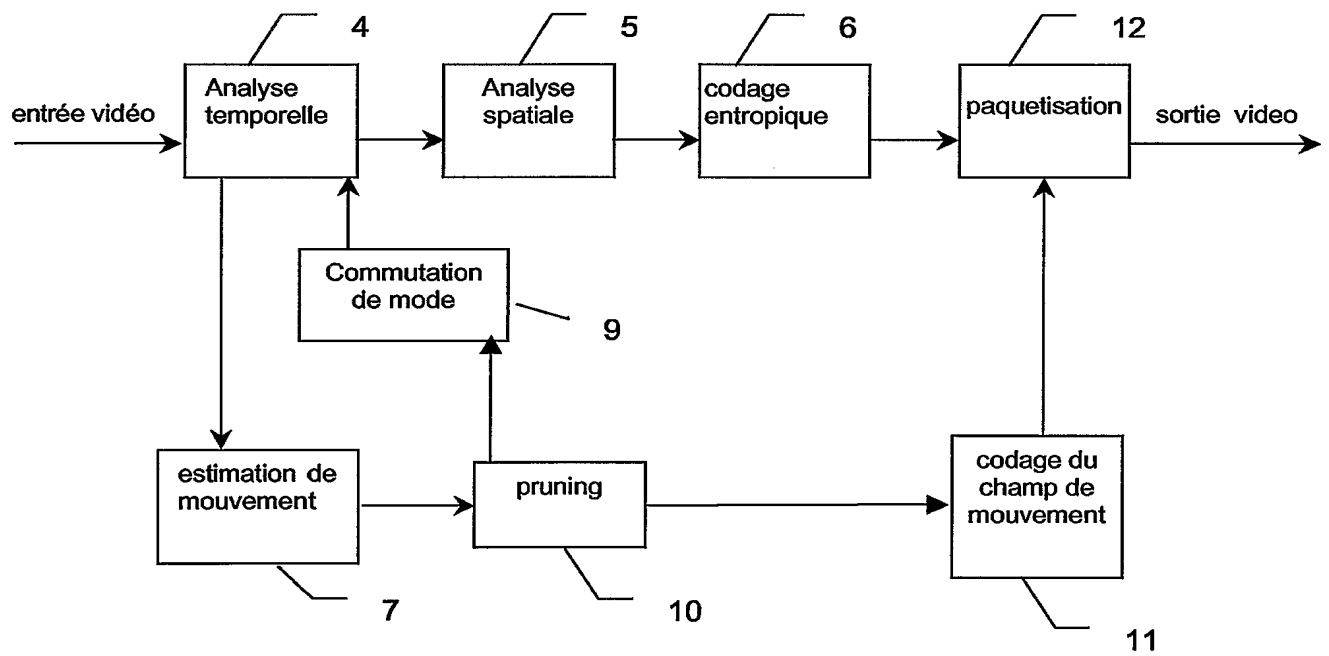


FIG.2

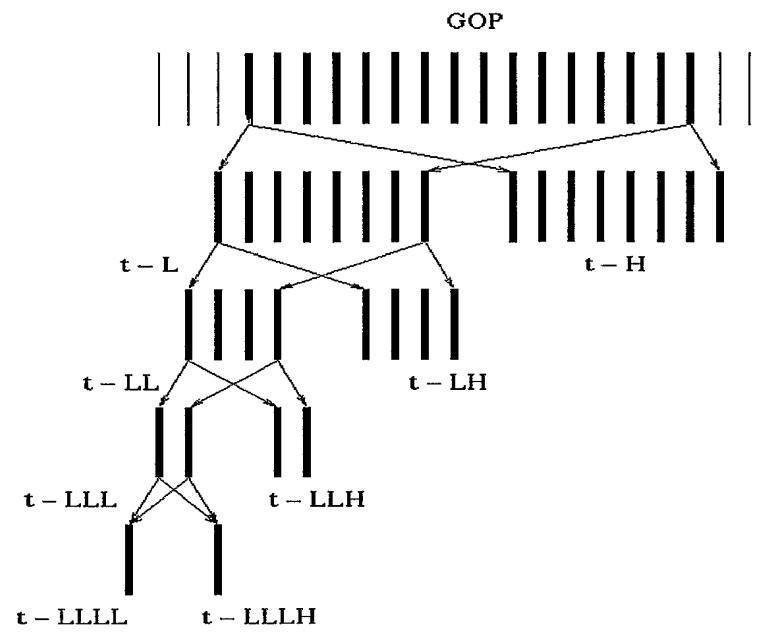


FIG 3

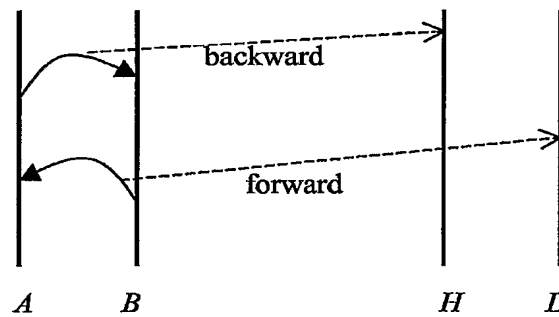


FIG 4

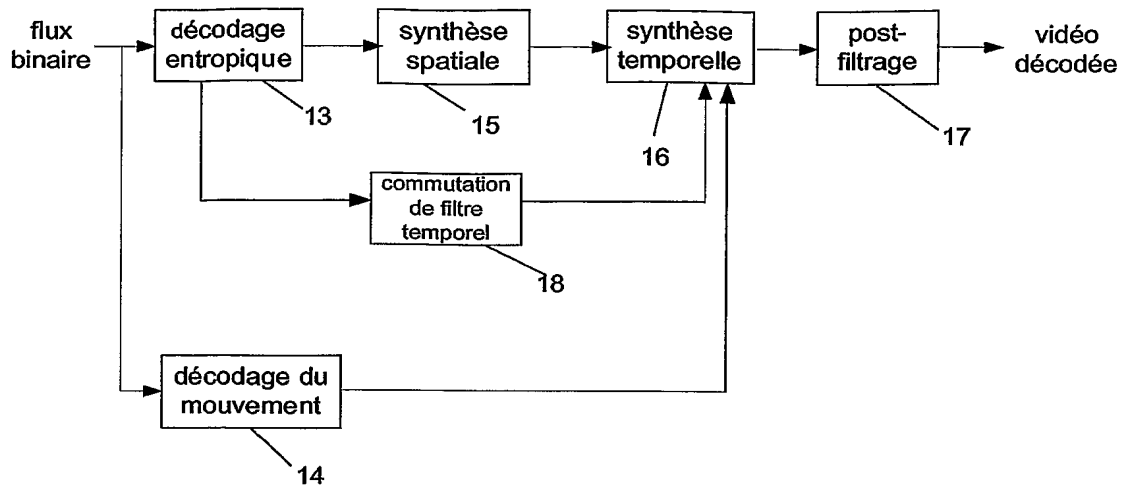


FIG.5

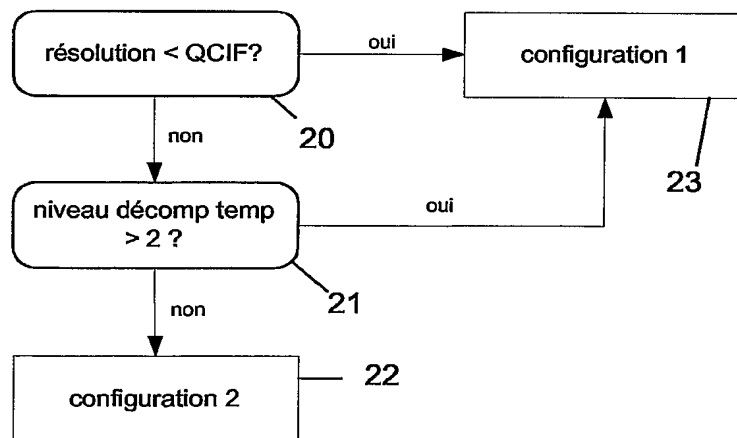


FIG.6

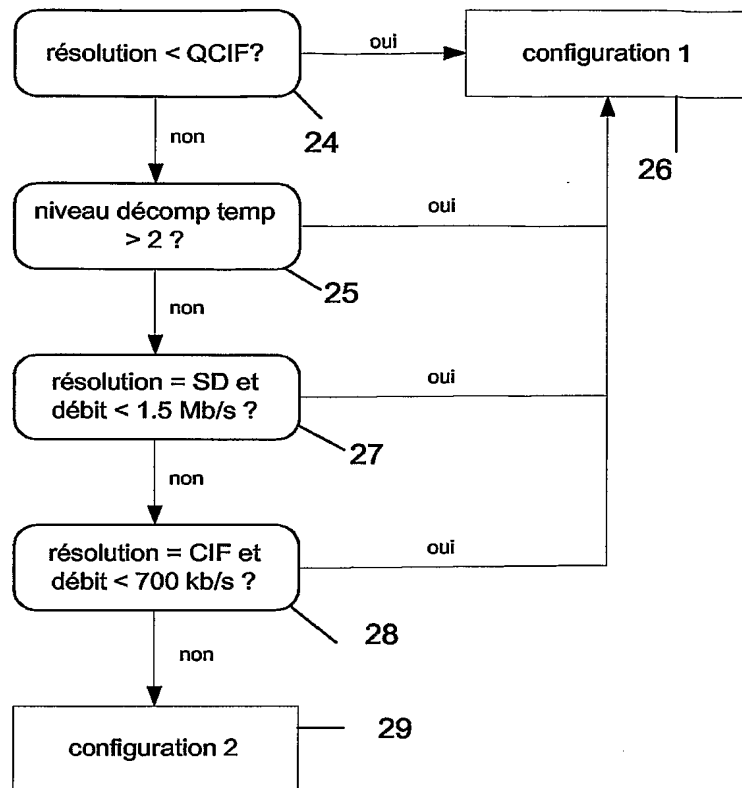


FIG.7

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/FR2005/050108

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 H04N7/26

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>TURAGA D S ET AL: "Content-adaptive filtering in the UMCTF framework"</p> <p>2003 IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING. PROCEEDINGS. (ICASSP). HONG KONG, APRIL 6 - 10, 2003, IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING (ICASSP), NEW YORK, NY : IEEE, US,</p> <p>vol. VOL. 1 OF 6,</p> <p>6 April 2003 (2003-04-06), pages III621-III624, XP010639149</p> <p>ISBN: 0-7803-7663-3</p> <p>the whole document</p> <p>-----</p> <p>-/-</p>	1-9

☒ Further documents are listed in the continuation of box C.☐ Patent family members are listed in annex.

## ° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&amp;" document member of the same patent family

Date of the actual completion of the international search

6 July 2005

Date of mailing of the international search report

25/07/2005

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
 NL - 2280 HV Rijswijk  
 Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
 Fax: (+31-70) 340-3016

Authorized officer

Georgiou, G

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/FR2005/050108

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>VAN DER SCHAAR M ET AL: "Unconstrained motion compensated temporal filtering (umctf) framework for wavelet video coding"</p> <p>2003 IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING. PROCEEDINGS. (ICASSP). HONG KONG, APRIL 6 - 10, 2003, IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING (ICASSP), NEW YORK, NY : IEEE, US,</p> <p>vol. VOL. 1 OF 6,</p> <p>6 April 2003 (2003-04-06), pages IIII81-III84, XP010639014</p> <p>ISBN: 0-7803-7663-3</p> <p>the whole document -----</p>	1-9

# RAPPORT DE RECHERCHE INTERNATIONALE

Demande Internationale No  
PCT/FR2005/050108

**A. CLASSEMENT DE L'OBJET DE LA DEMANDE**  
CIB 7 H04N7/26

Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB

**B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE**

Documentation minimale consultée (système de classification suivi des symboles de classement)  
CIB 7 H04N

Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche

Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si réalisable, termes de recherche utilisés)  
EPO-Internal

**C. DOCUMENTS CONSIDERES COMME PERTINENTS**

Catégorie °	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	<p>TURAGA D S ET AL: "Content-adaptive filtering in the UMCTF framework" 2003 IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING. PROCEEDINGS. (ICASSP). HONG KONG, APRIL 6 - 10, 2003, IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING (ICASSP), NEW YORK, NY : IEEE, US, vol. VOL. 1 OF 6, 6 avril 2003 (2003-04-06), pages III621-III624, XP010639149 ISBN: 0-7803-7663-3 le document en entier</p> <p style="text-align: center;">----- -/-</p>	1-9

☒ Voir la suite du cadre C pour la fin de la liste des documents

☐ Les documents de familles de brevets sont indiqués en annexe

° Catégories spéciales de documents cités:

"A" document définissant l'état général de la technique, non considéré comme particulièrement pertinent

"E" document antérieur, mais publié à la date de dépôt internationale ou après cette date

"L" document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée)

"O" document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens

"P" document publié avant la date de dépôt internationale, mais postérieurement à la date de priorité revendiquée

"T" document ultérieur publié après la date de dépôt internationale ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention

"X" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément

"Y" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier

"&" document qui fait partie de la même famille de brevets

Date à laquelle la recherche internationale a été effectivement achevée

6 juillet 2005

Date d'expédition du présent rapport de recherche internationale

25/07/2005

Nom et adresse postale de l'administration chargée de la recherche internationale  
Office Européen des Brevets, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Fonctionnaire autorisé

Georgiou, G



# RAPPORT DE RECHERCHE INTERNATIONALE

Demande Internationale No  
PCT/FR2005/050108

C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
A	<p>VAN DER SCHAAR M ET AL: "Unconstrained motion compensated temporal filtering (umctf) framework for wavelet video coding"</p> <p>2003 IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING. PROCEEDINGS. (ICASSP). HONG KONG, APRIL 6 - 10, 2003, IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING (ICASSP), NEW YORK, NY : IEEE, US,</p> <p>vol. VOL. 1 OF 6,</p> <p>6 avril 2003 (2003-04-06), pages IIII81-IIII84, XP010639014</p> <p>ISBN: 0-7803-7663-3</p> <p>le document en entier</p> <p>-----</p>	1-9